Plant Pathology (2011) Doi: 10.1111/j.1365-3059.2011.02491.x

Spatial analysis and incidence–density relationships for downy mildew on hop

D. H. Gent[a]*, J. L. Farnsworth[b] and D. A. Johnson[c]

[a]*US Department of Agriculture-Agricultural Research Service, Forage Seed and Cereal Research Unit, and Department of Botany and Plant Pathology, Oregon State University, Corvallis, OR 97331;* [b]*Department of Crop and Soil Science, Oregon State University, Corvallis, OR 97331; and* [c]*Department of Plant Pathology, Washington State University, Pullman, WA 99164-6430, USA*

The spatial pattern of downy mildew (*Pseudoperonospora humuli*) on hop (*Humulus lupulus*) was characterized over 4 years to aid in deriving an appropriate incidence–density relationship. From 472 disease assessments (datasets), discrete distributions were fitted to the datasets to determine aggregation of disease density. Where distributions were able to be fitted, the Poisson distribution fitted 4% of the datasets and the negative binomial distribution fitted 87% of the datasets. Larger-scale patterns of disease were assessed by autocorrelation and runs analysis; both indicated aggregation of diseased plants was less common than aggregation of disease within plants. Taylor's power law indicated disease density was aggregated and related to mean disease density in all years. Disease incidence and density were linked by saturation-type relationships based on the zero term of the negative binomial distribution or an empirical regression. Certain individual datasets were not described well by any incidence–density model, particularly when disease density was greater than about 0·8 diseased shoots per plant with the cultivar Cascade. When applied to 56 validation datasets, 88% of the variation in observed disease incidence was explained by the incidence–density models. Under conditions where sampling would be implemented for disease management, the requisite conditions appear to be in place for a binomial sampling plan for downy mildew.

Keywords: *Humulus lupulus*, *Pseudoperonospora humuli*, quantitative epidemiology

Introduction

Sampling is an essential component of informed disease management decisions. The level of a disease (measured as incidence, severity or density) is commonly estimated by sampling a field or other management unit. The specific purpose of estimating disease levels may be to assess efficacy of a management tactic applied previously, to determine if a fungicide application is warranted (Shoemaker & Lorbeer, 1977) or define the intensity of applications that should be made, or to provide an input to a disease forecasting model (Johnson & Coil, 1989).

Often, sampling can be simplified and expedited by simply noting the presence or absence of disease symptoms without quantifying the level of disease on individual sample units. Such approaches are referred to as binomial count sample plans. Binomial sampling plans can reduce sampling costs, particularly for pests that are difficult to enumerate, although this often comes at the expense of reduced precision (Binns *et al.*, 2000). The theoretical basis for any binomial sample plan is the functional relationship between mean disease density and incidence, which has been demonstrated for numerous plants diseases and arthropod pests (Jones, 1994; McRoberts *et al.*, 2003). In this context, disease density refers to the number (count) of lesions or other units of infection (e.g. diseased shoots) expressed relative to an entire plant or plant part such as a leaf, shoot or branch. Disease incidence refers to the number or proportion of plants or plant parts that are diseased. Binomial sampling allows one to estimate mean disease density per sampling unit based on the incidence of diseased individuals, although binomial sampling decreases precision and introduces bias into a sample compared to full-count sampling plans (Binns *et al.*, 2000). Development of a binomial sampling plan requires quantification of disease patterns to derive a suitable incidence–density relationship, as well as of disease density in multiple cultivars, fields and seasons to ensure that an incidence–density relationship is consistent and performs as intended during sampling plan validation.

*E-mail: gentd@onid.orst.edu

Disclaimer: The use of trade, firm, or corporation names in this publication is for the information and convenience of the reader. Such use does not constitute an official endorsement or approval by the United States Department of Agriculture or the Agricultural Research Service of any product or service to the exclusion of others that may be suitable.

Sampling is commonly conducted to assess the level of downy mildew (caused by *Pseudoperonospora humuli*) in hop (*Humulus lupulus*) production yards to obtain inputs needed for disease forecasting (Johnson & Coil, 1989) or other purposes. The disease may result in crop loss when cones, shoots or crowns become infected, and in the Pacific Northwest management tactics for the disease are applied most intensively during wet weather in spring and early summer to minimize shoot and crown infections (Skotland, 1961; Skotland & Johnson, 1983). Reductions in marketable yield vary tremendously depending on when and how severely the disease occurs, with the potential for complete crop loss in the worst case (Royle & Kremheller, 1981). Shoot infections are systemic and result in stunted, brittle shoots often referred to as 'basal spikes' that may support copious sporulation by the pathogen (Johnson & Skotland, 1985; Johnson *et al.*, 2009). The pathogen is readily disseminated by airborne sporangia, which are produced in a rhythmic pattern on relatively warm, humid evenings and dispersed in mid-morning (Royle & Kremheller, 1981; Johnson & Skotland, 1985). Because of the systemic nature of *P. humuli* in hop plants, the pathogen can also invade the root system and crown buds (Royle & Kremheller, 1981), resulting in persistent, perennial infections of hop plants (Coley-Smith, 1964). Consequently, the pathogen can be easily dispersed and established in hop yards via infected planting materials (Coley-Smith, 1965).

Growers and field scouts routinely estimate downy mildew severity by qualitative or quantitative assessments of the number of diseased shoots per plant. The incidence of plants with basal spikes is also an input to a forecasting model developed for downy mildew management in Washington State (Johnson & Coil, 1989). Assessing the density of diseased shoots is difficult because hop plants may produce hundreds of shoots and quantitative assessments of disease density are time-consuming and tedious. Disease assessments for downy mildew could be improved by binomial sampling where only the presence and absence of basal spikes would need to be determined, and disease incidence (proportion of plants with downy mildew) could be used to estimate mean disease density based on a model of the incidence–density relationship. To this end, the objectives of this study were to (i) characterize the spatial pattern of downy mildew on hop shoots within and among plants and (ii) derive an appropriate incidence–density relationship for use in later binomial sequential sampling.

Materials and methods

Field sites and data collection

The density of shoots with downy mildew was quantified in commercial hop yards in the Willamette Valley in western Oregon, USA during 2006–2009. All of the hop yards included in this study were located in Marion County in northwestern Oregon where most commercial hop production occurs in Oregon. Hops are long-lived perennial plants with annual shoots, and in the USA hops are grown under a permanent trellis approximately 5·5 m tall. Plant spacing can be in a regular lattice pattern or in wider rows with narrow spacing of plants within rows. In this study, plant spacing was either 2·1 × 2·1 m between plants and rows, respectively (narrow row spacing) or 1 × 4·25 m between plants and rows, respectively (wide row spacing). Yards planted using a narrow row spacing were irrigated by sprinklers, whereas yards planted on a wide row spacing were irrigated mostly by surface drip irrigation. The yards included in this study were planted to the cultivars Cascade (one yard), Centennial (one yard), Crystal (one yard), Liberty (one yard), Mt. Hood (one yard), Nugget (15 yards), Super Galena (one yard), Vanguard (one yard) and Willamette (seven yards). All yards were sampled in at least 2 years, except for one yard each of Nugget, Super Galena and Willamette that were sampled only during 2008. Resistance to downy mildew varied among the cultivars included in this study from highly susceptible (Crystal and Mt. Hood) to moderately resistant (Willamette; Johnson *et al.*, 2009). The age of the hop yards ranged from 2 to 20 years.

Model-construction dataset

A model-construction dataset was obtained from data collected during 2006–2008 from 24 commercial hop yards in which 472 disease assessments were conducted. Yards were assessed every 2–3 weeks beginning in early March and continuing through to early to mid-July. Disease density ($\hat{m}$) was assessed using a stratified sampling approach (Lohr, 1999). Stratified sampling was used instead of simple random sampling for several practical reasons. Stratified sampling in space ensures a representative sample over a target sampling area, in this case a given hop yard. Given that hop plants are widely spaced, stratified sampling (as described below) was more convenient and faster to execute than simple random sampling. Further motivation for stratified sampling was for consistency with sampling protocols for powdery mildew (caused by *Podosphaera macularis*), another important foliar disease of hop that growers routinely monitor by sampling (Turechek & Mahaffee, 2004).

To conduct the sampling, yards were stratified into multiple strata (H), where H = number of rows in a yard/20 (rounded-up to the nearest integer). A single number, r, between 1 and 20 was chosen randomly and the rth row from each stratum was sampled as a linear transect as described below. Herein, a row is considered synonymous with a transect. The stratum size was selected based on preliminary sampling (utilizing several sizes of strata) and practical considerations of other sampling protocols utilized for powdery mildew.

From each transect, the first 50 plants or 100 plants (N; sampling units) along the transect were inspected and the number of diseased shoots on each plant was recorded. Disease assessments were conducted by examining each plant for signs and symptoms characteristic of shoot infection by the downy mildew pathogen in the basal shoots, namely chlorosis and shortened internodes. This

involved moving leaves and shoots around near the base of each plant so that small shoots could be observed under the mat of basal foliage that often develops on hop plants. If downy mildew was presumptively identified, each basal shoot on the plant was inspected and considered diseased if sporulation was visible on the abaxial leaf surface under magnification with a × 50 hand lens. Over the course of the studies in Oregon nine different raters were involved with the disease assessments, a necessity when conducting such large-scale studies, but also a potential source of error in disease measurements. Raters were trained in disease assessment methods and their counts were compared to those of the first author, who was involved with most of the disease assessments, to ensure rater biases were minimized.

The number of sampling units assessed per stratrum was 50 in 2006 and 2007, but because of low disease incidence in the moderately resistant cv. Willamette, 100 sampling units were assessed in 2008. The total number of transects sampled in a given yard ranged from two to 10. With this sample size the proportion of the row sampled with a given transect varied among yards since the length of the rows varied among yards (ranging from 31 to over 400 plants). The precise length of every row was not recorded, although typically the number of plants sampled was about 25–50% of the total length of the row. In 6·5% of transects, the entire row was sampled.

Many factors may influence the observed pattern of plant diseases, including for example quadrat shape, pathogen dispersal characteristics, and edge effects (Xu & Ridout, 2000). In this analysis plants along the field edges were included in the spatial analysis because overall patterns of disease in the management unit (i.e. a hop yard) were of interest. In practice, growers sample both the edges and interiors of yards and make management decisions for the entire yard based on their appraisal of disease levels. Based on this practical consideration, this study sought to quantify overall patterns of downy mildew in commercial hop yards, including field edges.

Validation datasets

In 2009, 10 hop yards (five each of cvs Nugget and Willamette) were sampled every 2 weeks during early April to mid-July to obtain 80 validation datasets collected independently of the model-construction datasets. Sampling was conducted as described for the model-construction datasets. Because of the relatively low incidence of downy mildew in 2009, 100 plants were assessed per transect in the yards planted to cv. Willamette, as in 2008.

Additionally, 16 datasets collected previously (Johnson *et al.*, 1991) from hop yards in Washington State during 1988 and 1989 were used to validate the incidence–density relationships described below. A full description of these yards, including a detailed spatial analysis of the epidemics, is provided in Johnson *et al.* (1991). Briefly, four hop yards planted to the highly susceptible cv. Cluster L-1 located in the lower Yakima Valley or near the Yakima Indian Reservation were selected for downy mildew density assessments in 1988 and 1989. A rectangular area of plants (1·9–2·3 ha in size) was selected arbitrarily and the number of infected shoots on all plants in the sampling area was assessed at least twice during spring to early summer. Because the size and shape of the sample unit may affect some spatial analyses (Binns *et al.*, 2000), the datasets from Washington were subjected to simulated sampling. To do this, the sampling area in the Washington hop yards was stratified as described for the Oregon datasets and a transect was selected randomly from each stratum for sampling. These datasets were then used in the analysis described below.

Spatial analysis

Distributional analysis

The Poisson and negative binomial distributions were fitted to disease density data (i.e. the frequency distribution of the number of diseased shoots per plant) using the computer program DISCRETE (Gates & Ethridge, 1972). For discrete data with no upper limit, a good fit to the Poisson distribution is an indication of a random pattern of diseased plants, whereas a good fit to the negative binomial distribution is an indication of an aggregated disease pattern at the scale of the sampling unit (Binns *et al.*, 2000). Goodness of fit was determined with a chi-square test after pooling adjacent categories that had small expected frequencies until the cumulative frequency exceeded 1. Parameters were estimated using the maximum likelihood estimator in DISCRETE.

Autocorrelation

First- and second-order spatial autocorrelation statistics were calculated to quantify the similarity of disease density between neighbouring plants within individual transects (Madden *et al.*, 2007). Spatial autocorrelation analyses were not performed among plants in different transects because the location of these plants and transects was not recorded. From the 1936 rows where disease assessments were conducted, 349, 344 and 150 rows had at least one diseased shoot in 2006, 2007 and 2008, respectively. Count data were log-transformed before calculating autocorrelation statistics in MINITAB version 15 (Minitab Inc., State College, PA, USA).

Runs analysis

Ordinary and median runs analyses were performed to characterize larger-scale patterns of diseased plants within individual transects. For ordinary runs analysis, a plant was considered diseased if at least one diseased shoot was observed in that sampling unit. For median runs analysis, the median number of diseased shoots was calculated for each dataset, and plants were assigned a value of 1 or 0 if the number of diseased shoots on that plant was above or below the median, respectively. For both analyses, a run was defined as a succession of one or more plants with similar disease status (non-diseased or diseased; Madden *et al.*, 1982). The expected number of runs was calculated and used to produce a *Z*-statistic to test the null hypothesis that the number of runs was not

different significantly ($P \leq 0{\cdot}05$) from the expected number of runs, indicating a random pattern of disease among plants. Runs analysis were conducted using MINITAB version 15.

Power law analysis

For disease density or other discrete data with no upper limit, Taylor's power law (Taylor, 1961, 1984; Taylor & Taylor, 1977) provides a simple model to quantify aggregation at the scale of individual sampling units. Taylor's power law relates the observed variance (v_{obs}) to the population mean (m) for count data with no upper limit by

$$v_{\text{obs}} = am^b \tag{1}$$

where a and b are parameters estimated by regression after logarithmic transformation of Eqn 1 to

$$\ln(v_{\text{obs}}) = \ln(a) + b\log(m) \tag{2}$$

The parameter b is considered an index of aggregation, providing a measure of the degree of 'density dependence of aggregation' (Taylor & Taylor, 1977). For count data, $v_{obs} = m$ (i.e. $a = b = 1$), indicates a random pattern of disease density. Parameter estimates of a and $b > 1$ indicate aggregation at the scale of the sampling unit dependent on disease density.

Least squares regression was used to estimate the intercept and slope parameters of Taylor's power law using SAS version 9·2 (PROC REG, SAS Institute). Covariance analysis was conducted to test for the effects of the factors year of sampling (2006, 2007 and 2008), time of season (April, May, June, July), cultivar (Willamette, Nugget, and the remaining other cultivars), and row spacing ('narrow' or 'wide' row spacing) on the slope and intercept parameters of Taylor's power law. Covariance analyses were performed in SAS (PROC GENMOD) as described by Gent *et al.* (2008). Briefly, each of the factors was added individually as an intercept term and then as an interaction term with the slope. The analyses were conducted on each year, and then a separate analysis was conducted to determine the effect of year on estimates of ln (a) and b. A factor was considered significant if inclusion of the factor as a covariate significantly reduced the sum of square error (SSE) as compared to the null model without the factor. The significance level for the difference between SSE of the models was determined by an F test where F = (factor SSE/d.f. factor)/(model SSE/d.f. model), and d.f. = degrees of freedom.

Incidence–density relationships

Several approaches to characterize the incidence–density relationship between the incidence of plants with downy mildew and the density of diseased shoots were investigated. From the disease density data, disease incidence, $\hat{p}$, was calculated as the proportion of plants with at least one diseased shoot. Disease density data collected during 2006–2008 was used to fit the following incidence–density relationships, and the validation datasets collected in 2009 were used to the validate the models.

Poisson distribution

In the simplest case, an incidence–density was derived assuming that the number of infected shoots on plants followed a Poisson (random) distribution. In this case, the probability, p, of at least one diseased shoot being found on a sampling unit (plant) when mean disease density is m can be expressed as

$$p = 1 - P(0|m) = 1 - \exp(-m) \tag{3}$$

where $P(0)$ is the probability of plants with no diseased shoots. The probability p can be estimated directly from disease incidence from the sample of N plants. The inverse of Eqn 3 allows determination of m given an estimate of p by

$$m = -\ln(1 - p) \tag{4}$$

Negative binomial distribution with k *expressed by Taylor's power law*

Equations 3 and 4 may provide an adequate description of data if the spatial pattern of disease density is randomly distributed, although another distributional model may be needed when disease density is aggregated. Based on the results of the distributional analyses (described below), the negative binomial distribution provided a good description of most of the datasets. The negative binomial distribution is a flexible distribution with two parameters, the mean density m (sometimes referred to as μ) and an aggregation parameter k. For a fixed m, variance decreases with increasing k, and for very large values of k the negative binomial becomes indistinguishable from the Poisson distribution (Binns *et al.*, 2000). The zero term of the negative binomial distribution is $(m/k + 1)^{-k}$ and provides the number of sampling units with no diseased individuals. By taking the complement of the zero term, the proportion of sampling units with at least one diseased individual is given by

$$p = 1 - (m/k + 1)^{-k} \tag{5}$$

with the inverse of Eqn 5 being

$$m = k[(1 - p)^{-1/k} - 1] \tag{6}$$

The parameter k provides an indication of the aggregation of disease density at the scale of the sampling unit, although k may vary depending on the size (or shape) of the sampling unit and m (Binns *et al.*, 2000). When sampling is conducted in multiple fields, each field will have a separate value of k, although a single value of k is needed for designing a sequential sampling plan. Bliss & Owen (1958) described methods to calculate a common value of k, k_c, and provided diagnostic tests for determining the homogeneity of k values between datasets and the appropriateness of a k_c. In preliminary analyses, a regression of $1/k$ versus mean disease density had a significant and non-zero slope ($P < 0{\cdot}0001$) and intercept ($P < 0{\cdot}0001$), indicating the appropriateness of a k_c for these datasets was questionable (Bliss & Owen, 1958). Influence analysis (Quinn & Keough, 2002) identified five datasets as having high 'leverage', although the slope and intercept

terms in the regression still had significant ($P < 0{\cdot}0001$) and non-zero parameter estimates when these datasets were removed from the analysis. Therefore, a k_c from these datasets was not deemed appropriate.

For a given sampling unit size and shape, the density dependence of k also can be modelled empirically by Taylor's power law. Using the moment estimate of the aggregation parameter, Wilson & Room (1983) showed that one can express k in terms of m and a and b from Taylor's power law as

$$k = \frac{m^2}{am^b - m} \tag{7}$$

Substitution of k as defined in Eqn 7 into Eqn 5 leads to

$$p = 1 - \exp[-m(am^{b-1} - 1)^{-1} \ln(am^{b-1})] \tag{8}$$

However, an inverse of Eqn 8 does not exist and therefore this equation is not helpful for estimating m as a function of disease incidence and aggregation expressed by Taylor's power law parameters. Thus, other formulations of incidence–density relationships were explored as described below.

Empirical regression model

The Kono–Sugino equation also was used to estimate m from p (McRoberts *et al.*, 2003). In this approach, an empirical regression based on the complementary log-log transformation of disease incidence, $\mathrm{CLL}(p) = \ln[-\ln(1 - p)]$ was directly regressed on the natural log transformation of m through the equation

$$\mathrm{CLL}(p) = \ln(a') + b' \ln(m) \tag{9}$$

which corresponds to the relationship $p = 1 - \exp(-am^b)$. Parameters a' and b' in Eqn 9 are estimated by regression, and are distinct from parameters a and b in Taylor's power law. The inverse of Eqn 9 allows for estimation of m as a function of p and is expressed as

$$\ln(m) = \ln(\gamma) + \delta\,\mathrm{CLL}(p) \tag{10}$$

Equation 9 was fitted using the REG procedure in SAS. Predictions of m based on p based on Eqns 3, 8 and 10 were evaluated with the independent validation datasets.

Results

Disease density and incidence

From the 472 sampling events in the model-construction dataset, downy mildew was observed in 261 of the assessments. Disease density ranged from 0 to 24·39 diseased shoots per plant, with mean 0·52 and median 0·004 (Fig. 1a). Disease incidence ranged from 0 to 0·997, with mean 0·09 and median 0·004 (Fig. 1b).

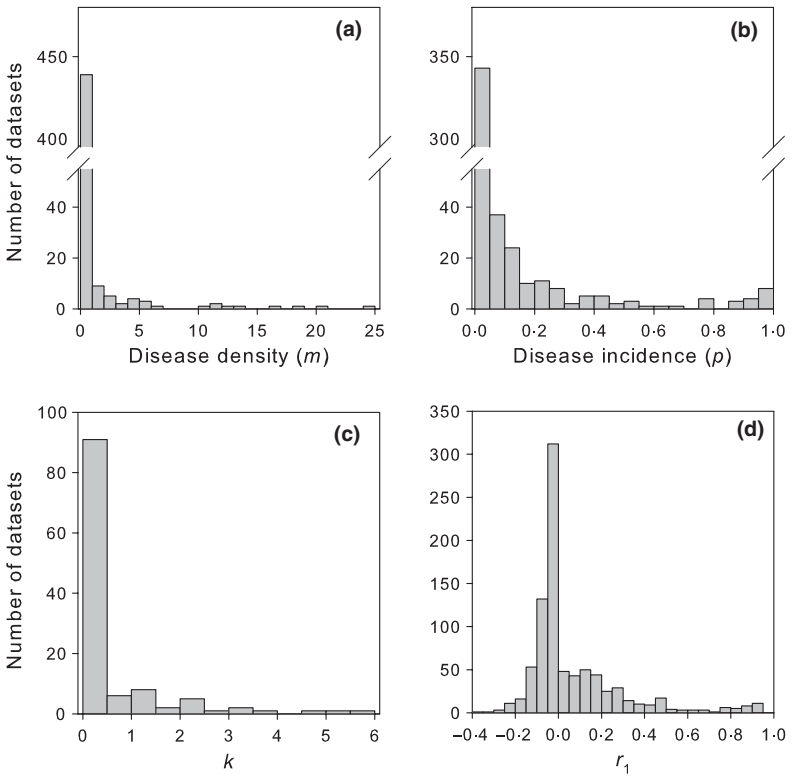

Figure 1 Frequency distribution of mean disease density m (a), disease incidence p (b), aggregation parameter k of the negative binomial distribution (c), and first-order autocorrelation statistic r_1 (d) for the density of shoots with downy mildew (*Pseudoperonospora humuli*) sampled from commercial hop yards in Oregon during 2006–2008. Bars are centered over binning values.

Spatial analysis

Distributional analysis

The maximum likelihood procedure in DISCRETE converged for 130 of the 261 datasets where $m > 0$ when fitting the Poisson distribution. In all cases when the procedure did not converge there were three or less diseased shoots present in the datasets. For the 130 datasets where the procedure converged, the Poisson distribution fitted only 4% of the datasets (Table 1). The negative binomial distribution provided an adequate fit in 87% of the 119 datasets where the maximum likelihood procedure converged for this distribution fitting, indicating a substantial degree of aggregation of diseased shoots within plants. The distribution of the aggregation parameter k was right-skewed with median 0·21 and mean 0·61 (Fig. 1c), and k tended to increase with m (Table 1).

Autocorrelation

For all row-level datasets where $m > 0$, the distribution of the first order autocorrelation statistic (r_1) was right-skewed with median −0·021 and mean 0·065 (Fig. 1d). The degree of autocorrelation indicated a low level of aggregation of disease density between plants when mean disease density was ≥ 1 diseased shoot per plant (Fig. 1d). Median values of r_1 when disease density was < 1 diseased shoot per plant ranged between −0·02 and 0·03, indicating that disease density was largely random between plants (Table 1).

Runs analysis

Ordinary runs analysis indicated that diseased plants were aggregated in 15% of the row-level datasets where disease was detected (Table 1). Aggregation of diseased plants did not appear to be systematically related to disease density, although median runs analysis suggested that the aggregation of downy mildew increased with mean disease density (Table 1).

Power law analysis

Taylor's power law was fitted to the 226 datasets where more than one diseased shoot was observed (Table 2; Fig. 2). The model provided a good fit to the data, with R^2 values of at least 0·95 for all years. Parameter estimates for b were > 1 in all years ($P < 0·0001$), whereas parameter estimates for ln (a) were > 0 in all years

Table 2 Slope and intercept parameter estimates of Taylor's power law fit to the density of hop shoots with downy mildew (*Pseudoperonospora humuli*) sampled from commercial hop yards in Oregon

Year	d.f.	b (SE)	ln (a) (SE)	R^2
2006	88	1·24 (0·03)	1·15 (0·08)	0·95
2007	105	1·26 (0·02)	1·32 (0·07)	0·96
2008	27	1·34 (0·06)	1·66 (0·20)	0·95
All	224	1·25 (0·02)	1·25 (0·05)	0·96

d.f.: degrees of freedom for regression; b and ln (a): slope and intercept estimates, respectively. SE: standard error of the mean.

Table 1 Frequency distribution models fitted to counts (density) of hop shoots with downy mildew (*Pseudoperonospora humuli*) and tests of aggregation for datasets collected from commercial hop yards in Oregon during 2006–2008

		Distribution[c]		Median values[d]		Runs analysis[e]	
Disease density class[a]	T[b]	Poisson	Negative binomial	k	r_1	Ordinary	Median
0	211	–	–	–	–	–	–
0·00–0·10	27	0·07 (27)	0·94 (16)	0·04	−0·02	0·15	0·14
0·10–0·50	52	0·06 (52)	0·92 (52)	0·12	−0·02	0·17	0·19
0·50–1·00	18	0 (18)	0·83 (18)	0·28	0·03	0·08	0·21
1·00–5·00	20	0 (20)	0·80 (20)	0·50	0·08	0·12	0·25
> 5·00	13	0 (13)	0·69 (13)	2·61	0·13	0·18	0·28
All (> 0)	130	0·04 (130)	0·87 (119)	0·21	−0·02	0·15	0·19

[a]Mean density of hop shoots with downy mildew per plant. Disease density classes end with the indicated value and start with the next highest value above that listed. Data are not presented for 131 datasets where three or fewer diseased shoots in total were present because of lack of convergence of the maximum likelihood estimation procedure used for distribution fitting. Disease density in these datasets ranged from 0·003 to 0·030 with mean = 0·015 and median = 0·009.

[b]Number of datasets (hop yards) in each disease density class where the maximum likelihood procedure converged when fitting the Poisson distribution.

[c]Proportion of datasets for which the said distribution provided an adequate description of the data as determined by a chi-square goodness-of-fit test ($P \leq 0·05$). The number of datasets in which the maximum likelihood estimation procedure converged is shown in brackets. Distributions were fitted using the program DISCRETE (Gates & Ethridge, 1972).

[d]Median estimated value of the negative binomial distribution parameter (k) and the first-order autocorrelation statistic (r_1). Autocorrelation was calculated with data for individual transects (rows) and not all transects within a yard. The number of row-level datasets for each disease density class was: 1055 (0), 372 (0·00–0·10), 259 (0·10–0·50), 84 (0·50–1·00), 99 (1·00–5·00), 67 (> 5·00) and 881 (all > 0).

[e]Proportion of datasets in which runs analysis indicated significant aggregation ($P \leq 0·05$). Runs analysis was calculated with data from individual transects and not all transects within a yard.

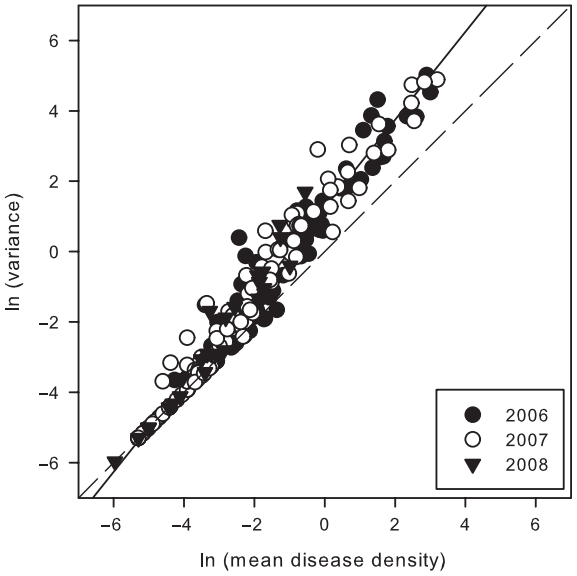

Figure 2 Relationship between the logarithms of mean disease density and variance of the density of downy mildew (*Pseudoperonospora humuli*) on hop shoots sampled from commercial yards in Oregon during 2006–2008. The solid line is the least squares regression fitted to the data; the dashed line is the line for a Poisson (random) distribution of disease density.

($P \leq 0{\cdot}0001$). Therefore, disease density was aggregated in all years, and the degree of aggregation was density dependent.

Covariance analysis indicated that ln (a) and b were affected by the time of season when disease assessments were conducted in 2006 and 2007 ($P \leq 0{\cdot}013$; Table 3). Row spacing also affected b in 2006 ($P = 0{\cdot}043$). The slope and intercept parameter estimates were similar between years ($P \geq 0{\cdot}07$; Table 4), and a regression based on the data pooled over years had ln (a) = 1·25 (SE = 0·05) and b = 1·25 (SE = 0·02).

Validation datasets

Spatial analysis

Among the 80 validation datasets collected in 2009 in Oregon, downy mildew density ranged from 0 to 2·70 diseased shoots per plant, with mean 0·18 and median 0·01. The incidence of plants with downy mildew ranged from 0 to 0·58, with mean 0·07 and median 0·01. A total of 50 datasets had at least one diseased shoot, and there were 40 datasets where more than one diseased shoot was observed. Among these 40 datasets, the maximum likelihood procedure for fitting the Poisson distribution in DISCRETE converged for 25 of the datasets; the Poisson distribution provided an adequate description of only one (4%) dataset. The procedure converged for 20 datasets when fitting the negative binomial distribution, and this distribution provided an adequate description of 18 (90%) of these datasets.

Taylor's power law was fitted to the 40 datasets from Oregon and 16 datasets from Washington where more than one diseased shoot was observed (Fig. 3). The

Table 3 Covariance analysis of the effect of cultivar, season and row spacing on intercept (ln [a]) and slope (b) parameters of Taylor's power law for the density of hop shoots with downy mildew (*Pseudoperonospora humuli*)

			ln (a)				b			
Factor and year[a]	d.f. model	d.f. factor	SSE	Diff.	F[b]	P	SSE	Diff.	F	P
2006										
Power law	88		29·87	–			29·87	–		
Cultivar	85	3	28·15	1·72	1·73	0·170	28·48	1·39	1·39	0·253
Season	86	2	21·97	7·90	15·47	0·000	26·30	3·57	5·83	0·004
Spacing	87	1	28·91	0·96	2·89	0·093	28·48	1·39	4·24	0·043
2007										
Power law	105		25·88	–			25·88	–		
Cultivar	102	3	23·48	2·40	3·48	0·019	24·08	1·80	2·54	0·061
Season	103	2	21·48	4·41	10·56	0·000	23·77	2·11	4·57	0·013
Spacing	104	1	25·70	0·19	0·76	0·386	25·88	0·01	0·02	0·875
2008										
Power law	27		5·27	–			5·27	–		
Cultivar	25	2	5·26	0·01	0·02	0·978	5·26	0·01	0·02	0·976
Season	25	2	5·08	0·20	0·49	0·619	5·12	0·15	0·36	0·698
Spacing	26	1	5·27	0·00	0·02	0·891	5·26	0·01	0·04	0·835

d.f. model: degrees of freedom for the model; d.f. factor: degrees of freedom for factor. SSE: sum of square error for the covariance model; Diff.: difference between the SSE of the binary power law model versus the binary power law model with each factor included in the analyses first as an intercept and then as a slope.

[a]Cultivars were categorized as 'Willamette', 'Nugget' or 'Other'. The factor 'season' refers to month of year and had three categories: April, May or June. The factor 'spacing' refers to plant arrangement with 3·5 m (narrow spacing) or 7 m (wide spacing) between rows.

[b]Significance level for the difference between SSE of the binary power law model versus binary power law model with each factor as determined by an F test, where F = (factor SSE/d.f. factor)/(model SSE/d.f. model).

Table 4 Covariance analysis of the effect of year of sampling on the intercept (ln [*a*]) and slope (*b*) parameters of Taylor's power law for the density of hop shoots with downy mildew (*Pseudoperonospora humuli*)

			ln (*a*)				*b*			
Factor	d.f. model	d.f. factor	SSE	Diff.	F^a	*P*	SSE	Diff.	*F*	*P*
Power law	224		63·01	–	–	–	63·01	–	–	–
Year	222	2	61·53	1·48	2·66	0·07	62·70	0·305	0·54	0·58

d.f. model: degrees of freedom for the model; d.f. factor: degrees of freedom for the factor 'year'; SSE: sum of square error for the covariance model; Diff.: difference between the SSE of the power law model versus the power law model with the factor 'year' included in the analyses first as an intercept and then as an interaction with the slope.

[a]Significance level for the difference between SSE of the power law model versus power law model with the factor 'year' as determined by an *F* test, where *F* = (factor SSE/d.f. factor)/(model SSE/d.f. model).

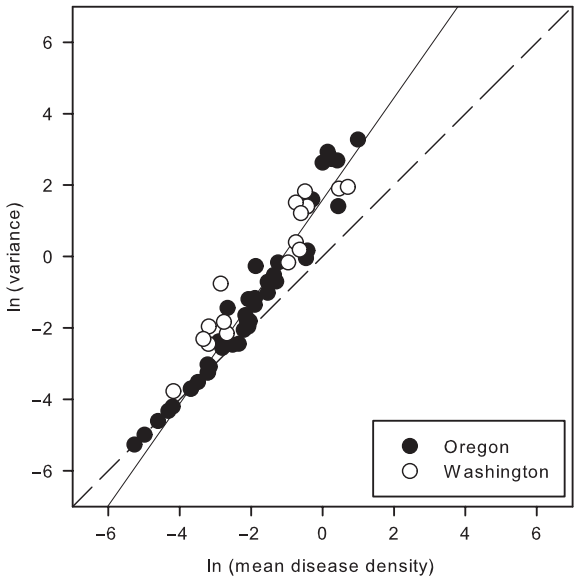

Figure 3 Relationship between the logarithms of mean disease density and variance of the density of downy mildew (*Pseudoperonospora humuli*) on hop in validation datasets collected from Oregon in 2009 and Washington in 1988 and 1989. The solid line is the least squares regression fitted to the data; the dashed line is the line for a Poisson (random) distribution of disease density.

estimated parameters of the regression were ln (*a*) = 1·60 (SE = 0·13) and *b* = 1·38 (SE = 0·05), with R^2 = 0·93. Parameter estimates for *a* and *b* were both significantly > 1 (P < 0·0001).

Incidence–density relationships

The incidence of plants with downy mildew was related to the mean density of shoots with downy mildew by a saturation-type relationship (Fig. 4a). Observed disease incidence increased more slowly with increasing mean disease density than expected under an assumption of a Poission distribution, as expected from aggregation of disease density. Incidence–density models based either on the negative binomial distribution or a complementary log-log transformation generally provided a reasonable fit to the data collected during 2006–2008, particularly when disease density was less than about 0·8. In the model based on the zero term of the negative binomial distribution, *p* increased more rapidly with *m* than in the empirical model based on the complementary log-log transformation. The logarithm of the mean number of diseased shoots per plant was well related to the complementary log-log transformation of disease incidence, with parameter estimates of ln (a') = −0·783 (SE = 0·028) and b' = 0·846 (SE = 0·009) and R^2 = 0·97 (P < 0·0001). The fit of this model was not improved by including the factor 'year' as an intercept term or interaction term with the slope ($P \geq 0{\cdot}059$).

Certain individual datasets were not described well by any incidence–density model, particularly when *m* was greater than about 0·8 diseased shoots per plant. For six of these datasets, *p* was substantially less than predicted based on the observed *m*; five of these datasets were collected from hop yards planted to cv. Cascade and one of cv. Nugget.

Validation dataset

When the number of diseased shoots per plant was assumed to follow a Poisson distribution, the predicted number of diseased shoots based on disease incidence generally was biased and greater than observed (Fig. 4b). Despite this bias, all three incidence–density models explained approximately 88% of the variation in observed disease incidence in the validation datasets based on a regression of predicted mean density on observed mean density. As observed in the model-development datasets, there was more variability in observed incidence–density relationship as *m* increased. Systematic prediction errors were not observed in the validation datasets for Oregon or Washington (Fig. 4b).

Discussion

Quantification of spatial patterns of plant disease is a prerequisite to designing statistically sound sampling methods, and for linking disease incidence and density (McRoberts *et al.*, 2003; Madden *et al.*, 2007). In this study, the incidence and density of downy mildew on hop shoots were quantified using a polyphasic approach to characterize patterns of the disease at multiple spatial

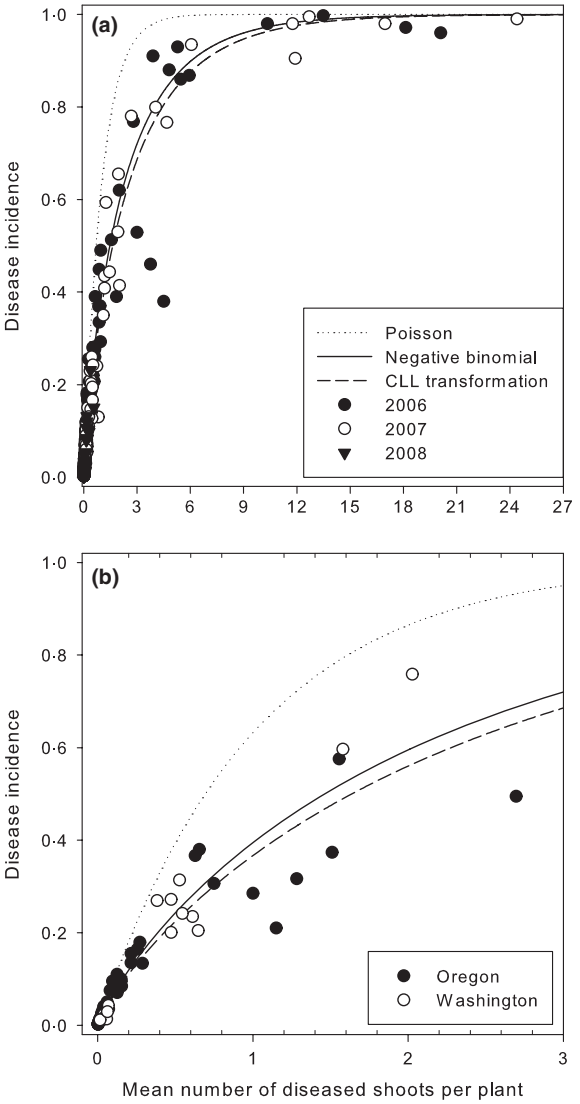

Figure 4 Observed and fitted relationships between the mean density of hop shoots with downy mildew (*Pseudoperonospora humuli*) and the incidence of hop plants with downy mildew used for model development (a) and validation with independent datasets collected in Oregon and Washington (b). CLL transformation refers to an empirical regression approach based on a complementary log-log transformation of disease density.

scales. At the scale of individual plants, the density of diseased shoots was highly aggregated, as indicated by the good fit of the negative binomial distribution and the poor fit of the Poisson distribution to most of the datasets. There were some datasets where neither distribution provided an adequate fit. This appeared to be mostly a computational issue related to convergence of the maximum likelihood procedure at low disease density, although there was a trend for the negative binomial to describe a smaller proportion of the datasets as m increased (Table 1). The increasing median value of the aggregation parameter k with increasing m suggests that the pattern of disease density became less aggregated, but not necessarily randomly distributed since the Poisson distribution did not fit any of these datasets. Other distributions also were fitted to the datasets, including the Neyman type A, Thomas double Poisson, Poisson with zeros, and logarithmic with zeros, but none was found to provide a better description than the negative binomial of the distribution of disease density at high values of m. Johnson *et al.* (1988, 1991) reported similar findings in their studies on hop downy mildew on highly susceptible cultivars in Washington State. In those studies, the negative binomial provided a good description of the data when p was relatively low, but no distribution provided an adequate fit when $p > 0{\cdot}16$ (Johnson *et al.*, 1991). In the current study, the negative binomial fitted 84% of datasets where $p > 0{\cdot}16$ and the goodness-of-fit of this distribution, as expected, was related more to the value m rather than p.

Power law analysis also suggested that disease density was aggregated at the scale of individual plants, and similarly indicated that the degree of aggregation was density dependent. Such a finding is typical of most plant diseases, as well as arthropod pests, at numerous spatial scales (Taylor, 1961, 1984; Taylor *et al.*, 1978; Xu & Madden, 2002; Li *et al.*, 2007; Madden *et al.*, 2007). Cultivar, time of the season when sampling was conducted, and row spacing influenced the intercept and slope parameters in certain years, but no factor had a consistent effect on the intercept or slope during all 3 years. Parameter estimates also tended to be relatively stable among all years (Table 4).

Beyond the scale of individual plants, autocorrelation and median runs analyses suggested that larger-scale patterns of disease density among plants were less common within rows. There was a trend for a greater proportion of datasets to display aggregation of disease density as m increased, but overall aggregation of disease density appeared to be most distinct at the scale of individual plants with the cultivars evaluated in the current study. In other spatial analyses of downy mildew diseases, epidemics tend to dominated by spatial aggregation at small spatial scales (e.g. Madden *et al.*, 1995; Scott *et al.*, 2003) but also at very large spatial scales (Wu *et al.*, 2001; Ojiambo & Holmes, 2011). The similar patterns at both scales may be largely influenced by the high fecundity and dispersal characteristics of downy mildew pathogens. Epidemics of downy mildew on hop appear to be an extreme case of small-scale aggregation, presumably owing to the systemic nature of rootstock infection and persistence on individual plants (Coley-Smith, 1964). Although large-scale patterns of disease were not considered in the current study, one might expect that *P. humuli* could be dispersed over distances similar to those of the very closely related pathogen *P. cubensis* (Choi *et al.*, 2005; Runge *et al.*, 2010; Mitchell *et al.*, 2011). Dispersal of *P. cubensis* is thought to occur over the range of hundreds of kilometres, with evidence of spatial autocorrelation apparent at ranges of up to 1000 km (Ojiambo & Holmes, 2011). Apparently, no studies have investigated long distance dispersal of *P. humuli* or landscape-level patterns of hop downy mildew.

It is relevant to contrast this finding with previous studies conducted with hop cultivars susceptible to the crown rot phase of downy mildew (Johnson *et al.*, 1991). *Pseudoperonospora humuli* can invade and persist in the root system of hop plants (Coley-Smith, 1964, 1965; Royle & Kremheller, 1981), resulting in a root and crown rot in certain susceptible cultivars (Johnson *et al.*, 2009). In Washington State, severe outbreaks of downy mildew occur on average in about one-third of years (Johnson *et al.*, 1983). The hop production regions in central Washington are semiarid and outbreaks of downy mildew tend to occur only during seasons with frequent spring rains. Consequently, some cultivars that are highly susceptible to crown rot can be produced commercially since downy mildew occurs sporadically in this semiarid environment. Conversely, downy mildew is endemic in the cool, maritime climate of western Oregon and only cultivars with a high degree of tolerance to the crown rot phase of the disease can be produced successfully (Skotland & Johnson, 1983). In Washington State, Johnson *et al.* (1988, 1991) reported that when disease incidence was relatively low downy mildew was aggregated primarily at the scale of individual plants, and larger-scale patterns of disease were related to the age of the hop yard because of the death and replanting of plants affected by downy mildew. In contrast, plant death and subsequent replanting is uncommon in Oregon since the cultivars produced are tolerant to crown rot. Consequently, in the current study, aggregation of downy mildew was most conspicuous at the scale of individual plants, and larger-scale patterns of disease were related to disease density rather than age of the yard.

Indeed, the relationship between incidence and density of downy mildew on shoots was well described by a model based on the zero term of the negative binomial distribution, with a few notable exceptions. The relationship between incidence and density of downy mildew appeared to deviate in individual fields from the model predictions when disease density was relatively high ($m > 3$). This deviation could be caused, in part, by the negative binomial not being the appropriate distribution for describing disease density when disease density is relatively high. As stated previously, the proportion of datasets described by the negative binomial decreased with increasing disease density. However, an incidence–density relationship based on an empirical regression did not improve the fit of the incidence–density model. Part of the variability observed appears to be related to inherent differences in cv. Cascade, for which estimates of m based on p tended to be underestimated. In the USA, cv. Cascade is relatively tolerant to downy mildew crown rot, but susceptible to shoot infection (Johnson *et al.*, 2009). The mechanism of this tolerance may affect the dynamics of disease incidence and density differently than in other tolerant cultivars included in this study, such as Willamette.

Regardless of the biological mechanisms involved in generation of the observed incidence–density relationships, deviation in the relationship at relatively larger values of m probably are of little practical consequence for routine disease management. Although critical densities associated with economic thresholds have not been derived for downy mildew, growers routinely manage the disease at levels much less than three diseased shoots per plant (Johnson & Coil, 1989). In practice, some growers may make downy mildew fungicide applications preventatively, at the first detection of disease, or when disease incidence or density begins to increase based on sampling. In all of these cases, control measures are applied relatively early in disease outbreaks, e.g. when $m < 1$. The relationship between p and m was approximately linear when $m < 1$, which is where sampling for downy mildew generally would be of greatest practical value for determining the need for control measures.

A prerequisite for use of a binomial sampling plan is a stable, or at least predictable, relationship between disease incidence and density over space and time. Although this relationship seems to be less predictable with cv. Cascade, overall the incidence–density relationships based on the negative binomial distribution or an empirical regression model performed adequately during validation with three cultivars. It is interesting to note that the incidence–density relationship seemed to hold even for the datasets collected in Washington State with the highly susceptible cv. Cluster (Fig. 4b). As stated previously, this cultivar is susceptible to crown rot and no probability distribution (including the negative binomial) was found to describe the distribution of diseased shoots on this cultivar when $p > 0{\cdot}16$ (Johnson *et al.*, 1991). Under relatively low levels of disease density considered in this study, approximately 88% of the observed variability in disease incidence could be explained by the incidence–density models. Therefore, under conditions where sampling would be most valuable for disease management, the requisite conditions appear to be in place for development and use of a binomial sampling plan for hop downy mildew. Such a binomial sampling plan will be developed in future research.

Acknowledgements

Financial support was provided by USDA-ARS CRIS 5358-21000-035-00, Washington State University, and USDA-CSREES Western IPM Center. Bill Turechek provided helpful assistance with the distributional analysis described in this manuscript. We thank Bill Pfender and Larry Madden for their review and comments on an earlier draft of the manuscript. We also thank the many individuals who assisted in disease assessments and the hop growers who allowed us access to their farms for sampling.

References

Binns MR, Nyrop JP, van der Werf W, 2000. *Sampling and Monitoring in Crop Protection: the Theoretical Basis for Designing Practical Decision Guides*. Wallingford, UK: CABI Publishing.

Bliss CI, Owen ARG, 1958. Negative binomial distributions with a common *k*. *Biometrika* **45**, 37–58.
Choi YJ, Hong SB, Shin HD, 2005. A re-consideration of *Pseudoperonospora cubensis* and *P. humuli* based on molecular and morphological data. *Mycological Research* **109**, 841–8.
Coley-Smith JR, 1964. Persistence and identification of downy mildew *Pseudoperonospora humuli* (Miy + Tak) Wilson in hop rootstocks. *Annals of Applied Biology* **53**, 129–32.
Coley-Smith JR, 1965. Infection of hop rootstocks by downy mildew *Pseudoperonospora humuli* (Miy. et Tak.) Wilson and its control by early-season dusts. *Annals of Applied Biology* **56**, 381–8.
Gates CE, Ethridge FG, 1972. A generalized set of discrete frequency distributions with FORTRAN program. *Mathematical Geology* **4**, 1–24.
Gent DH, Turechek WW, Mahaffee WF, 2008. Spatial and temporal stability of the estimated parameters of the binary power law. *Phytopathology* **98**, 1107–17.
Johnson DA, Coil K, 1989. *Hop Downy Mildew Program. A Program for Personal Computers*. Pullman, WA, USA: Washington State University Bulletin, MCP0008.
Johnson DA, Skotland CB, 1985. Effects of temperature and relative humidity on sporangium production of *Pseudoperonospora humuli* on hop. *Phytopathology* **75**, 127–9.
Johnson DA, Skotland CB, Alldredge JR, 1983. Weather factors affecting downy mildew epidemics of hops in the Yakima Valley of Washington. *Phytopathology* **73**, 489–93.
Johnson DA, Alldredge JR, Fisher R, 1988. Spatial pattern of downy mildew in hop yards. *Phytopathology* **78**, 1389–95.
Johnson DA, Alldredge JR, Allen JR, Allwine R, 1991. Spatial pattern of downy mildew in hop yards during severe and mild disease epidemics. *Phytopathology* **81**, 1369–74.
Johnson DA, Engelhard B, Gent DH, 2009. Downy mildew. In: Mahaffee WF, Pethybridge SJ, Gent DH, eds. *Compendium of Hop Diseases and Pests*. St Paul, MN, USA: APS Press, 18–22.
Jones VP, 1994. Sequential estimation and classification procedures for binomial counts. In: Pedigo LP, Buntin GD, eds. *Handbook of Sampling Methods for Arthropods in Agriculture*. Boca Raton, FL, USA: CRC Press, 175–205.
Li B-H, Yang J-R, Li B-D, Xu X-M, 2007. Incidence-density relationship of pear scab (*Venturia nashicola*) on fruits and leaves. *Plant Pathology* **56**, 120–7.
Lohr SL, 1999. *Sampling: Design and Analysis*. Pacific Grove, CA, USA: Brooks/Cole Publishing.
Madden LV, Louie R, Abt JJ, Knoke JK, 1982. Evaluation of tests for randomness of infected plants. *Phytopathology* **72**, 195–8.
Madden LV, Hughes G, Ellis MA, 1995. Spatial heterogeneity of the incidence of grape downy mildew. *Phytopathology* **85**, 269–75.
Madden LV, Hughes G, van den Bosch F, 2007. *The Study of Plant Disease Epidemics*. St Paul, MN, USA: APS Press.
McRoberts N, Hughes G, Madden LV, 2003. The theoretical basis and practical application of relationships between different disease intensity measurements in plants. *Annals of Applied Biology* **142**, 191–211.
Mitchell MM, Ocamb CM, Grünwald NJ, Mancino LE, Gent DH, 2011. Genetic and pathogenic relatedness of *Pseudoperonospora cubensis* and *P. humuli*. *Phytopathology* (in press), doi: 10.1094/PHYTO-10-10-0270.
Ojiambo PS, Holmes GJ, 2011. Spatiotemporal spread of cucurbit downy mildew in the eastern United States. *Phytopathology* **101**, 451–61.
Quinn GP, Keough MJ, 2002. *Experimental Design and Data Analysis for Biologists*. Cambridge, UK: University Press.
Royle DJ, Kremheller HT, 1981. Downy mildew of the hop. In: Spencer DM, ed. *The Downy Mildews*. New York, NY, USA: Academic Press, 395–419.
Runge F, Choi YJ, Thines M, 2010. Phylogenetic investigations in the genus *Pseudoperonospora* reveal overlooked species and cryptic diversity in the *P. cubensis* species cluster. *European Journal of Plant Pathology* **129**, 135–46.
Scott JB, Hay FS, Wilson CR, Cotterill PJ, Fist AJ, 2003. Spatiotemporal analysis of epiphytotics of downy mildew of oilseed poppy in Tasmania, Australia. *Phytopathology* **93**, 752–7.
Shoemaker PB, Lorbeer JW, 1977. Timing initial fungicide application to control botrytis leaf blight epidemics on onions. *Phytopathology* **67**, 409–14.
Skotland CB, 1961. Infection of hop crowns and roots by *Pseudoperonospora humuli* and its relation to crown and root rot and overwintering of the pathogen. *Phytopathology* **51**, 241–4.
Skotland CB, Johnson DA, 1983. Control of downy mildew of hops. *Plant Disease* **67**, 1183–5.
Taylor LR, 1961. Aggregation, variance, and the mean. *Nature* **189**, 732–5.
Taylor LR, 1984. Assessing and interpreting the spatial distributions of insect populations. *Annual Review of Entomology* **29**, 321–57.
Taylor LR, Taylor RAJ, 1977. Aggregation, migration and population mechanics. *Nature* **265**, 415–21.
Taylor LR, Woiwod IP, Perry JN, 1978. The density-dependence of spatial behaviour and the rarity of randomness. *Journal of Animal Ecology* **47**, 383–406.
Turechek WW, Mahaffee WF, 2004. Spatial pattern analysis of hop powdery mildew in the Pacific Northwest: implications for sampling. *Phytopathology* **94**, 1116–28.
Wilson LT, Room PM, 1983. Clumping patterns of fruit and arthropods in cotton, with implications for binomial sampling. *Environmental Entomology* **12**, 50–4.
Wu BM, van Bruggen AHC, Subbarao KV, Pennings GGH, 2001. Spatial analysis of lettuce downy mildew using geostatistics and geographic information systems. *Phytopathology* **91**, 134–42.
Xu XM, Madden LV, 2002. Incidence and density relationships of powdery mildew on apple. *Phytopathology* **92**, 1005–14.
Xu XM, Ridout MS, 2000. Effects of quadrat size and shape, initial epidemic conditions, and spore dispersal gradient on spatial statistics of plant disease epidemics. *Phytopathology* **90**, 738–50.